NAYAK
EXHIBIT O

Meta
Documentation ▾   Trust & Safety ▾   Community ▾   Try Llama ↗   Download models   🔍

Community Stories

Discover the possibilities of building on Llama

Medical

PeacePal AI

PeacePal AI is a mental health companion app featuring engaging conversations with multiple unique AI companions that learn in real-time with realistic voice generation, provide goal tracking & support, guided meditations, and much more! With 25K+ Instagram followers, 2M+ animation views, and a passionate team of young entrepreneurs, we're making mental wellness solutions accessible and affordable for everyone worldwide!

› Learn more

Medical

Patsnap

The continuously evolving biomedical brain provides biopharmaceutical companies with comprehensive acceleration and efficiency improvements across R&D decision-making, data insights, knowledge management, and clinical trials, forming a one-stop solution.

› Learn more

Design

Ubitus

Ubitus leverages locally trained LLMs, including Llama 70B and 405B, to create AI avatars that enable multilingual communication, personalized support, and dynamic engagement. These avatars function as bilingual teaching assistants, AI VTubers like Ubi-chan, corporate brand representatives, and immersive robotics, transforming interactions in education, entertainment, and enterprise.

› Learn more

Tech

ELYZA

ELYZA is a leading company in large language model (LLM) development in Japan. Since 2019, we've been investing in LLM research and development. We're promoting the social implementation of LLMs in various fields such as finance and human resources through support for LLM utilization, development of custom LLMs, and the creation and provision of AI SaaS solutions

› Learn more

Tech

Qevlar

Qevlar augments SOC teams with advanced AI capabilities by leveraging large language models (LLMs), especially Meta's Llama. By enabling AI-driven investigations, Qevlar reduces response times, improves accuracy, and ensures security across complex workflows.

› Learn more

Consumer

Convirza

Leading AI company Convirza, a software platform designed to evaluate agents via AI models, helps customers analyze millions of hours of calls to produce actionable insights fast with conversation analytics. By leveraging Predibase's multi-LoRA serving infrastructure, Convirza achieved a 10x reduction in operational costs with Llama compared to OpenAI, an 8% improvement in F1 scores, and increased throughput by 80%, allowing them to efficiently serve over 60 performance indicators across thousands of customer interactions daily.

› Learn more

View more use cases   —

Tech

Neuron

On-Device AI Coding powered by Llama

Neuron revolutionizes AI coding with quantized Code Llama 7B and Llama 2 7B models running fully on-device. Boost productivity with instant responses, safeguard sensitive data, and optimize resources—no cloud needed. Seamlessly integrate Neuron into your workflow via a local server, connecting effortlessly to VS Code for an efficient and secure development experience.

> Learn more

Tech

Vecflow AI

Agentic AI for legal work

Vecflow designs AI agents for legal research, contract review and drafting. Their flagship product, Oliver, connects with internal systems to deliver insights and automate repetitive workflows. Legal organizations demand uncompromising standards in both data security and accuracy – that's why Vecflow built Oliver on Llama 3.1 405B. With Vecflow, you get deep legal domain expertise backed by enterprise-grade security infrastructure.

> Learn more

Education

Manda

AI tutor helps secondary students in the UK with math and English

Llama 3-powered Manda brings affordable AI tutoring to students anywhere in the UK. Online tutoring company TLC created Manda by training a Llama 3 model on over 9,100 hours of anonymized tutoring session transcripts from fully qualified teachers.

> Learn more

Impact

Dubai Holding

Llama Design Drive

Dubai Holding, a globally diversified investment company, participated in the Llama Design Drive, an AI accelerator powered by Meta. The program brought together startups and corporates to drive innovation using Meta's Llama 3.1 model. Dubai-based startup CarbonSifr won two challenges focused on improving greenhouse gas emissions reporting.

> Learn more

Impact

Nextpeer

Nora: AI Mental Health Companion by Nextpeer built on Llama

Nextpeer's Nora platform is transforming mental health support by utilizing Llama's advanced language capabilities to provide AI-driven companionship 24/7. Nora delivers personalized guidance, multimodal support, and secure, anonymous conversations. Built on Llama 3.2 and pre-trained on specialized mental health datasets, Nora offers empathetic understanding, contextual awareness, and continuous improvement through user feedback.

> Learn more

Tech

Patronus AI

Introducing Lynx, the leading hallucination detection model.

Patronus AI is pushing the boundaries of AI innovation with their Lynx model, a cutting-edge hallucination detection solution built on fine-tuned Llama 3 70B model. With unparalleled 87.4% accuracy and multilingual support, Lynx outperforms industry alternatives at catching hard-to-detect mistakes. Lynx empowers enterprises across healthcare, finance, and retail to combat inaccuracies and deliver reliable outcomes.

> Learn more

Tech

Hunar AI

Making Frontline hiring fast, easy and predictable

Hunar AI uses Llama 3.1 to streamline frontline & deskless workforce recruitment by efficiently sorting candidate profiles using structured data and matching them to job search parameters on their platform. It also uses Llama for their multilingual candidate assessments, facilitating the second stage of the recruitment process, while all candidate communication is seamlessly managed through WhatsApp Business Suite.

> Learn more

Medical

MMed-Llama 3

MMed-LLama 3

MMed-LLama 3 leverages the performance and diverse pretraining of Llama-3-8B to create a multilingual language model for healthcare applications. By introducing a large scale Multilingual Medical Corpus (MMC) and domain-specific pretraining, the specifically tailored model excels in medical question-answering with rationale across 21 fields and six languages.

> Learn more

Tech

CodeGPT

How Llama helped CodeGPT become one of the top AI-powered coding assistants

CodeGPT leverages Llama to boost developer productivity by 30%, streamlining tasks such as debugging, code generation, and onboarding. By integrating Llama, CodeGPT enables AI powered code suggestions, autonomous project folder generation, and expert AI agents for APIs and frameworks. This innovative approach transforms coding workflows, making development more intuitive and efficient for

Tech

WiggleBot AI

Your AI-powered hybrid work assistant

Learn more about WiggleDesk and how they leverage Llama to transform hybrid work management. By integrating Llama 3.1 70B into their platform, WiggleDesk's smart office assistant, WiggleBot AI, streamlines desk booking, schedule management and provides personalized support. WiggleDesk helps manage busy offices at organizations like The Duke of Edinburgh's Award, NHS and Local Councils. By utilizing Meta's Large Language Models with their custom data

Consumer

Untukmu AI

How online gifting site is using Llama to protect customer privacy

Untukmu.AI, an Indonesian platform that offers personalized gift recommendations for birthdays, anniversaries, and corporate holidays, has developed a unique approach to protecting the privacy of their customers' data, leveraging Llama on edge. The team designed a semi-decentralized personal assistant that ensures the company won't have access to customer data.

developers worldwide.

> Learn more

pipeline, WiggleDesk achieves advanced security, scalability, and reliability which enhances user experience, boosts productivity and simplifies complex workflows.

> Learn more

> Learn more

Impact

Envision

Perceive Possibility

Envision, a pioneering assistive technology company, leverages Llama 3.1 advanced AI capabilities to empower individuals who are blind or have low vision. By integrating Llama's language processing and computer vision, Envision's innovative products - including a mobile app and prototype on Meta's Project Aria glasses - can articulate visual information into speech, enabling users to access everyday information independently.

> Learn more

Tech

SNCF

Llama 3.1 in Action: Multilingual Virtual Assistant at Paris "Gare du Nord"Description:

"Show the trains for Paris". Explore how Sncf Group, Alp Valley and their partners integrated Llama 3.1 to create a voice-enabled virtual assistant for Paris' Gare du Nord, the largest station in Europe, unveiled at Vivatech Paris and the Paris 2024 Summer Olympics. This technical showcase demonstrates advanced module development for multilingual support and efficient workflow validation and real-time inference.

> Learn more

Tech

torchchat

Bringing Llama models to edge devices using PyTorch

The open-source community uses PyTorch libraries to fine-tune and run Llama models on various devices. Torchchat brings Llama to edge devices, preserving privacy and eliminating network connectivity needs. It leverages ExeruTorch, AOTInductor, quantization, and export to accelerate inference and enable broader platform support. This integration unlocks novel, privacy-aware experiences built on Lama on edge devices, utilizing PyTorch's widely-used deep learning capabilities.

> Learn more

Design, Education, Tech

Caddy AI

Introducing Caddy AI

Experience immersive storytelling with Caddy, a cutting-edge Meta AI powered platform for design, education and collaboration enabling an "IronMan" experience. With Caddy, users can import CAD style 3D models of simple and complicated objects alike and interact with the AI assistant on Quest to seamlessly visualize, analyze and build objects. It's how 3D design was always meant to be! Caddy is built on Llama 3.1 and seamless-m4T-v2 models for multilingual input and real time translation across players.

> Learn more

Education

Upeo Labs

How one Kenyan startup is working to solve local challenges with Llama

Kenya-based Upeo Labs is a generative AI research and development startup aiming to solve local challenges. Its initial offering, Somo-GPT, is a support tool for high school students, teachers, and parents. Powered by Llama 3 and Llama 3.1, Somo-GPT serves as a multi-subject teaching assistant, where people can ask questions on subjects like menstrual health, mental health, biology, geography, physics, history, and more, while interacting with a conversational chat interface.

> Learn more

Tech

Aive

Automating Video summarization and smart clipping with Llama

Aive, an innovative AI platform for video post-production, harnesses the power of Llama 3 to automate the creation of video summaries and smart clippings. By analyzing long-form content such as interviews, shows, or commercials, Aive quickly identifies moments and condenses them into shorter, engaging formats optimized for different platforms like Facebook and Instagram. This cutting-edge integration enables brands such as LVMH, L'Oréal, and Stellantis to streamline their video workflows, reduce costs, and maximize content impact across channels, all within a secure and scalable AI environment.

> Learn more

Gaming

Latitude

The Future of AI-generated games

Latitude, the developer of AI Dungeon, a popular role-playing game, recently migrated to OctoAI's optimized Llama 3.1 models (8B and 70B), achieving significant performance gains, including improved output quality, increased efficiency, and substantially lower costs. The transition from Llama 3.0 was completed in just half a day, with user blind tests verifying a 10-15% improvement in output quality, while token consumption increased by triple digits. Overall the migration from proprietary models to open source models like Llama has decreased their costs by five times.

> Learn more

Impact

Pratham Education Foundation

AI-Powered Chatbot Revolutionizing Early Childhood Care and Education

Managing early childhood education and health becomes challenging for a parent or caregiver in resource poor settings with limited access to information on these issues and subjects. For this reason, the Pratham Shah PraDigi Innovation Center has developed 'BaalSakhi', a chatbot powered by Meta's

Llama 2 (70B), which integrates seamlessly via WhatsApp, and offers personalized and on-demand support for early childhood care and education. BaalSakhi helps parents and caregivers through their child's early development journey, drawing insights from trusted sources such as government departments, UNICEF, WHO, Pratham, pediatric doctors, and childcare experts.

> Learn more

Business

KPMG LLP

Leveraging Llama to solve business needs

KPMG, an Audit, Tax and Advisory professional services firm, has leveraged Llama for multiple use cases with clients across industries. Through a tailored GenAI solution, The firm is helping the wholesale credit team at a major U.S. bank to explore secure, open-source LLMs, empowering the client to review and interrogate complex loan applications more quickly and efficiently. These efforts are positioning the team to take existing automation to the next level—dramatically improving the speed and accuracy of document review and processing.

> Learn more

Education

Cognitive Learner

A series of Open Source Bilingual LLMs for Indic Languages (Kannada-English)

Cognitive Labs is building Amabri, a large language model created to adapt LLMs to diverse Indian languages, starting with Kannada and English, and to develop efficient training processes. The model was trained on a modest 1 billion-token dataset, aiming to bridge the linguistic gap and observe world knowledge acquisition in LLMs.

> Learn more

Tech

TuneAI

Tune AI: Helping enterprises deploy fine-tuned Llamas trained on their data and infrastructure

Tune AI is a fine-tuning and deployment platform that assists large enterprises with custom use-cases. For an enterprise leader in the information services space, Tune AI selected Llama 3 in the interest of data security and privacy due to it being open source, to index a massive 7B+ page digital library in the academia and government division to bring down costs from manually indexing each document at $18 in an hour to $0.60 per document within 7 minutes, achieving a 99% accuracy rate.

> Learn more

Business

Clearpoint

Automating a refactor of over 500 code files using Llama 3

ClearPoint integrated Llama 3 model with Google's code scanner lint to perform an automated refactor in a large Kotlin code base. With less than 500 lines of automation code, we produced a migration that runs with a single build command without human intervention. Automating this category of code change would not have been possible without language models and it is satisfying to produce it in a way that does not risk intellectual property or vendor lock-in by employing local models like Llama 3. We estimate performing the code change manually would take 11 engineering weeks.

> Learn more

Business

Melooha

Harnessing Llama 3 70B: Revolutionizing Vedic Astrology with AI-Driven Insights

In the evolving landscape of technology, the fusion of ancient wisdom with modern computing has opened new frontiers. We have developed a hyper-personalized Vedic astrological engine, which is a blend of ancient knowledge and modern computing. Designed to deliver hyper-personalized, "credible", astrological insights, our engine harnesses the strengths of a rich astrological corpus and in house built prediction engine. At higher level there are two critical components in astrological engine: "Insight Detailing" and "Engaging Communication". The first component, driven by the Meta model Llama 3 70B, significantly enhances the accuracy and determinism of the astrological insights generated. This model's robust capabilities enable deep analysis, allowing for a profound exploration of astrological complexities.

> Learn more

Financial Business

Persado

Transforming Financial Services Marketing with Generative AI

Persado's Motivation AI, which utilizes advanced ML, NLP, and deep learning transformer models, creates emotion-informed messages that are 96% more effective in driving action. The LLM layer of Persado's platform is built upon the highly trusted and audited open-source Meta Llama collection of models, leveraging their capabilities for processing and generating natural language based on deep learning algorithms. The Llama models have been fine-tuned to an instruction model that serves the purpose of generating effective and on-brand marketing communication. The model has been refined with a proprietary dataset comprising hundreds of thousands of multilingual marketing content generation instructions.

> Learn more

Impact

Jacaranda Health

Empowering mothers to seek care at the right place and time

Jacaranda Health's PROMPTS program leverages AI and machine learning to support maternal health in Africa. By integrating Meta Llama, the platform efficiently triages incoming questions, identifies urgent cases, and provides critical support to expecting mothers in Kenya. Next, PROMPTS aims to enhance its capabilities by incorporating additional risk factors and conversational patterns to better monitor and support at-risk individuals, improving health outcomes through data-driven insights in order to build a comprehensive health navigator that will ultimately reach 1.3 million women across Sub-Saharan Africa.

> Learn more

Consumer

Shorthills AI

Healthcare

OpenBioLLM

Tech

NinjaTech AI

BestViewReviews Upgrades to Meta Llama 3 for Enhanced Product Insights

BestViewReviews uses GenAI and LLMs to analyze millions of customer reviews and expert blogs, providing unbiased and independent advice on products. Initially powered by Llama 2 7B, they've upgraded to Llama 3 8B, delivering even more precise and insightful information to empower customers in making better-informed buying decisions.

→ **Learn more**

Breakthrough in Medical AI: Introducing Llama3-OpenBioLLM-70B & 8B

Saama's AI Research Lab has achieved a significant milestone in healthcare AI innovation. After three years of rigorous research and collaboration with medical experts, they have collected high-quality datasets and developed the most openly available medical-domain LLMs to date. Research Led by Ankit Pal (https://aadityaura.github.io/), the Llama3-OpenBioLLM-70B & 8B models have outperformed industry giants in the biomedical domain, setting a new state-of-the-art for models of their size.

→ **Learn more**

Meta Llama 3 Powers NinjaTech AI's Next-Gen Agents

NinjaTech AI harnesses the strength of Llama 3, leveraging AWS's Trainium & Inferentia2 chips to unlock the 70B model's full potential. This powerful combination enables the development of cutting-edge AI agents, driving innovation in reinforcement learning. With Llama 3 and PyTorch at its core, NinjaTech AI's agentic models (Ninja LLMs) achieve significant gains in efficiency, scalability, and cost savings, revolutionizing the future of AI development.

→ **Learn more**

Tech

Copy.ai

CopyAI Leverages Meta Llama to Streamline Content Workflows

CopyAI, a pioneering GTM AI platform, leverages Llama on Lamini's LLM platform to accelerate content generation for go-to-market efforts. To tackle a complex content categorization challenge for a Fortune 100 client, involving over 900 categories, CopyAI fine-tuned a high-precision classifier based on Llama, served in production on Lamini. By combining Llama's capabilities with Lamini's high performance platform, CopyAI successfully automated content categorization, streamlining processes and driving efficiency.

→ **Learn more**

Healthcare

Transcripta Bio

From Genes to Medicine — Orchestrating Gene Expression for Therapeutic Results

Transcripta Bio harnesses Meta Llama 3 and BioNeMo for accelerated intelligent drug discovery. Its proprietary artificial intelligence modeling suite, Conductor AI, uses its Drug-Gene Atlas to help discover and predict the effects of new drugs at transcriptome scale.

→ **Learn more**

Education

Arivihan

Democratizing Education with Meta Llama

Arivihan, India's 1st 100% automated online learning platform, uses Llama to generate text responses and personalized lecture scripts, making high-quality education extremely affordable. With 82% accuracy, Llama has helped answer over 1,00,000 queries and check 6,000+ subjective answers, enabling Arivihan to democratize education while maintaining affordability and scalability.

→ **Learn more**

Impact

NoHarm.ai

NoHarm.ai uses Meta Llama to improve hospital discharge summaries for transition of care

NoHarm.ai, a healthcare non-profit startup in Brazil, makes it easier for the healthcare system to organize and communicate patients' information about their hospitalization, all in a data secure way. NoHarm Discharge Summary is an artificial intelligence, open-source web application that extracts the most important information from inpatients' records with Named Entity Recognition system (NER) then uses Llama 2 to write a discharge summary in Portuguese for physicians to validate.

→ **Learn more**

Agriculture

KissanAI

A Multilingual AI Agriculture Assistant

KissanAI, released Dhenu Llama 3, a cutting-edge AI model tailored for farmers in India. The fine-tuned model on Meta's Llama 3 8B architecture, has enhanced capabilities and superior benchmarks for agriculture and farming use-cases. The assistant can understand voice and text inputs and has also been integrated with platforms like WhatsApp making it accessible to farmers across India. The assistant currently is tailored for 22 languages including 9 Indian languages.

→ **Learn more**

Education

Mathpresso

MathGPT: Leveraging Llama 2 to create a platform for highly personalized learning

Mathpresso's QANDA platform makes personalized learning accessible globally. With Llama 2, they created MathGPT, an accurate math-specific LLM. Open-source models enable flexible, domain-specific educational products, democratizing access to quality education. Mathpresso and Upstage showcase the impact of open-source AI on education and innovation.

→ **Learn more**

Tech

Barcelona Supercomputing Center

Aloe: A Family of Fine-tuned Open Healthcare LLMs

Aloe is a healthcare-specific LLM, built on Meta Llama 3, with advanced training and inference mechanisms. It boasts +10 accuracy points and sets a new standard for ethical performance in healthcare LLMs, using policy alignment and Direct Preference Optimization. Aloe achieves state-of-the-art results for open healthcare 7B LLMs.

→ **Learn more**

Legal
Blinder

Blinder: Using Llama to equip law firms and compliance professionals with AI redaction tools and training

Blinder, a startup based in North Carolina uses Meta Llama to anonymize attorney privileged data while making it faster and easier to generate legal documents. They are also using Llama to track terms of service, privacy policies, costs, and data policies, alerting users to changes and recommending new models and apps based on usage data.

› Learn more

Impact
Hound

A web app harnessing AI to combat human trafficking

Hound is a web app harnessing AI to combat human trafficking. It was built using Meta Llama 3 and RAG to generate visual clues from survivor testimony, bridging information gaps to combat human trafficking. The app generates visual clues from survivor testimony, bridging information gaps to help law enforcement investigate and dismantle trafficking infrastructure.

› Learn more

Tech
NetSafe

Netsafe Leverages Meta Llama for Robust Redaction Tool

Netsafe, a not-for-profit, uses Meta Llama models to develop a robust redaction tool for harmful digital communications. The tool utilizes a tuned LLM model and the AI4Privacy dataset to automatically redact sensitive information, enabling faster and more effective harm resolution. Netsafe's data pipeline, built on Dagster, Postgres, and BigQuery, sources, redacts, and aggregates data from various channels, minimizing victim impact and augmenting Digital Harm Resolution Officers' skills.

› Learn more

Tech
ML Commons

ML Commons Adopts Meta Llama 2 70B for the MLPerf Inference Benchmark

ML Commons has incorporated Meta's powerful and accessible Llama 2 70B parameter model into version 4.0 of its MLPerf Inference benchmark. The published results demonstrate the performance potential of a range of platforms for running one of the most demanding and capable Large Language Models.

› Learn more

Tech
UC Berkeley

UC Berkeley Introduces RAFT: Enhancing Domain-specific RAG with Azure AI and Meta Llama 2

Discover how UC Berkeley researchers used Meta Llama 2 to create Retrieval-Augmented Fine-Tuning (RAFT), a novel approach that improves domain adaptation in language models like Llama 2. Explore RAFT's benefits, implementation using Llama 2 on Azure AI Studio, and the team's experience working with this versatile and adaptable model.

› Learn more

Education
FoondaMate

FoondaMate: AI Study Assistant Built on Meta Llama 2 Supports Over 3 Million Students in Achieving Academic Success

FoondaMate is an AI-powered study assistant built on Llama 2 that caters to over 3 million students on WhatsApp seeking exam and homework support. It provides instant, round-the-clock assistance to help save time, conserve data, and support learning. By leveraging the capabilities of Meta's Large Language Model, FoondaMate offers a convenient and efficient tool for students looking to optimize their learning experience.

› Learn more

Nonprofit
Odia Generative AI

Advancing AI for Indic Languages with OdiaGenAI

Odia Generative AI is an open-source project and community with a mission to expand the capabilities of LLMs to get regional language context, especially for Odia. Odia is one of the most widely spoken languages in India (40 million speakers). OdiaGenAI developed and released Odia Llama, Llama 2-fine tuned LLM for the Odia language. They are also exploring building applications on top of this model like AI chatbots, AI tutors, and more.

› Learn more

Tech
Taiwan LLM

Taiwan LLM Revolutionizes Traditional Chinese Language Processing with Cultural Context and Advanced Pre-Training Built on Meta Llama 2

Taiwan LLM is a pioneering Large Language Model that specifically caters to the Traditional Chinese language, with a focus on the variant used in Taiwan. Leveraging a comprehensive pre-training corpus and instruction fine-tuning datasets, Yen-Ting Lin, a PhD candidate at National Taiwan University, has developed a model that not only understands the complexities of Traditional Chinese but also embodies the cultural context of Taiwan.

Tech
Elyza

Elyza Leverages Meta Llama 2 to Develop Japanese Language Model for Advanced NLP Applications

Elyza, a Japanese startup focused on AI and natural language processing, has taken advantage of Llama 2 to develop a Japanese Large Language Model (LLM). This language model is designed to process and understand Japanese text, enabling applications such as language translation, chatbots, and text generation.

› Learn more

Healthcare
Mendel

Revolutionizing Healthcare AI: Mendel's Hypercube Platform Integrates Meta Llama 2 Based LLM for Natural Language Processing

Mendel created Hypercube, an AI platform using a hypergraph semantic layer on clinical data for applications like trial matching, National Comprehensive Cancer Network (NCCN), and patient cohorting. The proprietary Eloquent language allows querying and retrieving information from the hypergraph. To bridge the gap for users unfamiliar with Eloquent, Mendel's AI team used Llama 2 to convert natural language into Eloquent.

Impact

Refik Anadol Studio: Large Nature Model

Fusing art, technology, and nature for environmental advocacy

Artist Refik Anadol's Large Nature Model, an upcoming open-source generative AI model dedicated to nature, is a sensory experience emphasizing the interconnectedness of ecosystems. The model, which tapped into open-source data from institutions like the Smithsonian Institution and London's Natural History Museum, aims to promote awareness about environmental concerns and stimulate inventive solutions using generative AI.

Llama 2 is used to generate vibrant nature scenes based on rainforest data, with waterfalls, forests, animals, birds, and flowers. The installation includes custom generative sound and scents, allowing interaction with the raw dataset, helping to make the idea of environmental nature more real.

⟩ Learn more

Photo credit: Refik Anadol Studio

Tech

Lamini

Lamini is the LLM development for enterprises to create and control their own LLMs

Lamini is the LLM development platform for enterprises to create and control their own large language models (LLM) like Llama 2. The platform enables software development teams to create efficient LLMs with proprietary data with minimal effort, using only a few lines of code. Additionally, Lamini gives users the flexibility to deploy it securely either on-premise or in the cloud while also allowing for easy scaling of LLM inference and training from a single GPU to thousands.

⟩ Learn more

Tech

Sarvam

Advancing AI for Indic languages with OpenHathi and Llama 2

Sarvam AI is developing Llama 2 extensions for Indic languages and released its first Hindi LLM leveraging Llama 2 in partnership with AI4bharat. Hindi LLM is trained under compute and data constraints to demonstrate GPT-3.5-like performance on Indic languages with a frugal budget.

⟩ Learn more

Tech

Upstage

Upstage's Solar LLM Outperformed top models in Open-Source Language Model Leaderboard

Upstage develops a range of AI-powered solutions for businesses. Their large language model (LLM), Solar 70B, fine-tuned on Llama 2, secured the top position on the Open LLM Leaderboard when it debuted, a true testament to the opportunities that are unlocked by open innovation.

⟩ Learn more

Tech

OctoAI

Llama Guard for AI Trust and Safety

Generative AI platform OctoAI is using Llama Guard to detect unsafe traffic amid billions of customer inferences. This additional layer of safety flags policy violations that slip past existing moderation and filters, contributing to a safer LLM experience for customers and their users.

⟩ Learn more

Gaming

Together.AI

Infinite Craft & Together.AI

Neal Agarwal developed Infinite Craft using Llama 2 70B, allowing users to create new items by combining existing elements, while safely avoiding bad results with Llama Guard. The AI enabled surprisingly logical but witty results, making the game delightful to play, and helping it quickly go viral. Leveraging Together.AI for high performance inference, Neal was able to easily and elastically scale the game to meet the huge demand of over 650,000 tokens per second.

⟩ Learn more

Academic

LiGHT@Yale/EPFL

EPFL and Yale team up to make and validate a new Large Language Model to Democratize access to Medical Knowledge

Researchers at EPFL's School of Computer and Communication Sciences and Yale School of Medicine have developed Meditron, an open-source large language model trained on carefully curated medical guidelines to improve access to evidence-based information for clinical decision-making. Meditron outperforms all other open-source models on a range of medical benchmarks and nears the performance of much larger closed models like GPT4. They are now evaluating its performance in the real world with a MOOVE [Massive Online Open Validation and Evaluation] led by volunteer doctors across various specialties.

› Learn more

Business

LyRise

Unlocking AI for businesses

An Egypt-based talent-matching start-up using a chatbot built on Llama that interacts like a human recruiter connecting companies with AI-skilled job seekers across Africa, helping businesses find and hire top AI and data talent from a pool of high-quality profiles across various industries.

› Learn more

Healthcare

Mayo Clinic

RadOnc-GPT: A large language model for radiation oncology

Mayo Clinic fine-tuned Llama 2 to create RadOnc-GPT, a specialized Large Language Model (LLM) tailored for radiation oncology. This innovative application of Llama 2 boasts improved clarity, specificity, and clinical relevance with a localized, data-secure model.

› Learn more

Gaming

Niantic

Virtual pets in Peridot are now more lifelike (and adorable) than ever using generative AI

Niantic, the creator of Pokémon GO, launched a new feature in their AR-first game Peridot, which uses Llama 2 to generate environment-specific reactions and animations based on what the pet characters are interacting with and seeing in the real-world.

› Learn more

Technology

Zoom

Meet Zoom AI Companion, your new AI assistant

Zoom introduced AI Companion, a generative AI assistant that leverages a federated approach incorporating Zoom's own LLM built on Llama 2, as well as third-party models. AI Companion is designed to enhance productivity and collaboration by helping compose chat and email messages, summarizing meetings, and more, all while enabling data privacy and control over AI capabilities.

› Learn more

Business

Gupshup.io

#1 Conversation Cloud for customer acquisition, marketing, commerce and support

Their flagship CX Product Suite 'Converse' (spanning Click to Chat Ads, Campaign Manager, Bot Studio and Agent Assist) is now powered by ACE LLM which is a family of domain-specific models, fine-tuned for enterprise-grade conversations based on Llama 2 models.

› Learn more

Pharmaceutical

Nanome

Nanome launches AI copilot MARA, a scientific AI copilot

Molecular design XR startup, Nanome, launches MARA to aid scientists in biopharma R&D by orchestrating scientific workflows, executing routine cheminformatics tasks, and offering informative responses to scientific queries.

› Learn more

Share your Meta Llama story

Are you using Meta Llama in a unique or innovative way? We want to hear from you! Submit your story and share how you're using Meta Llama to make a difference.

› Go to form

Developer Sessions

Explore the latest developments in AI, mixed reality and wearables from sessions at Connect 2024.

Watch now

Meta Connect 2024

Our approach

Responsibility

Responsible Use Guide

To promote a responsible, collaborative AI innovation ecosystem, we've established a range of resources for all who use Llama 2: individuals, creators, developers, researchers, academics, and businesses of any size. The Responsible Use Guide is a resource for developers that provides recommended best practices and considerations for building products powered by LLMs.

› Responsible Use Guide

Open approach

Our investment in open trust and safety

With Meta Llama Guard we are furthering our commitment to an open approach to build generative AI with trust and safety in mind. We believe in transparency, open science, and cross-collaboration, and to date, we've released over a thousand open-source libraries, models, datasets, and more.

› Learn more

Ready to start building with Meta Llama?

Visit our getting started guide for resources to help you set up Meta Llama including how to access the model, hosting, how-to and integration guides, and more.

Get started guide

Stay up-to-date

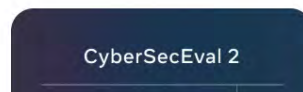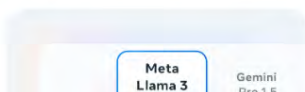### Our latest updates delivered to your inbox

Subscribe to our newsletter to keep up with the latest Meta Llama updates, releases and more.

› Learn more

Resources

Explore more on Meta Llama

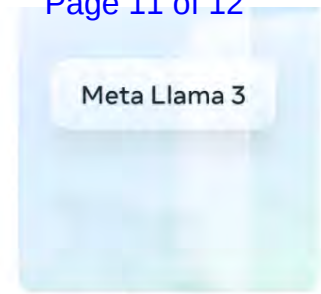| MMLU 5-shot | 82.0 | 81.9 |
GPQA 0-shot	39.5	41.5 CoT
HumanEval 0-shot	81.7	71.9
GSM-8K 8-shot, CoT	93.0	91.7 11-shot

Introducing Meta Llama 3: The most capable openly available LLM to date

Our responsible approach to Meta AI and Meta Llama 3

CYBERSECEVAL 2: A Wide-Ranging Cybersecurity Evaluation Suite for Large Language Models

Meta Llama 3 repository

> Read the blog

> Read the blog

> Read the paper

> View repository

Share your Llama story

First name *		Last name *
Company name *		Link to your website or social platform of choice
Email *		Phone number
Category *		Country / Region: United States of America
On what platform is this being used *		

Please briefly describe your project *

Which llama model was used? *

☐ Meta Llama 3

☐ Meta Llama 2

☐ Meta Code Llama

☐ Meta Llama Guard 2

By submitting this form, I agree to having Meta add my story to the Community Stories webpage.

For more information about how Meta handles your data please read our Meta Data Policy.

Accept and continue

∞ Meta

Documentation
Overview
Models
Getting the Models
Running Llama
How-To Guides
Integration Guides
Community Support

Trust & Safety

Community
Community Stories
Open Innovation AI Research Community
Llama Impact Grants

Resources
AI at Meta Blog
Meta Newsroom
FAQ
Privacy Policy
Terms
Cookies

Overview
Responsible Use Guide